Additional File 2

Figure S1 The null distribution of node tightness *S* depends on the number of leaves. This dependence is illustrated for all the benchmarks and dissimilarity – linkage combinations analyzed. In each case the distributions of *S* are shown for nodes with 2, 5 and 20 leaves. Each plot is based on 5000 randomizations of the respective data set.

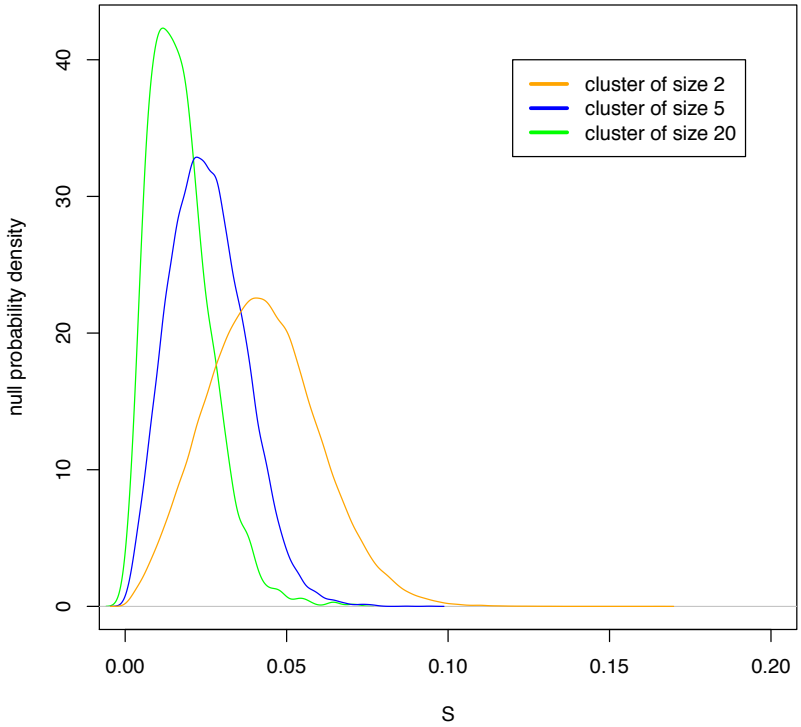
Links to S1

Figure S2 Empirical *p*-value estimates for tightness compared to EVT-based estimates. Combined results for all tree nodes in all benchmark studies are shown. For each benchmark the combinations of dissimilarity and linkage are enumerated in the same order as they appear in Table 2. Displayed are the values corrected for hypothesis multiplicity (*cf* the Methods section). Empirical estimates are based on $1000 \times N$ randomizations each, N being the number of leaves. EVT estimates are based on 1000 randomizations each. If the empirical p-value estimate based on these 1000 randomization is large, the EVT algorithm defaults to this estimate. The corresponding points are shown by empty symbols of the appropriate shape and color. The diagonal dashed line indicates the identity. The vertical dashed line indicates the minimal multiplicity-corected empirical p-value $[1 - (1 - p_e)^{N-2}] / (n_r + 2)$, where N is the number of leaves and n_r is the number of randomizatons.

FigureS1

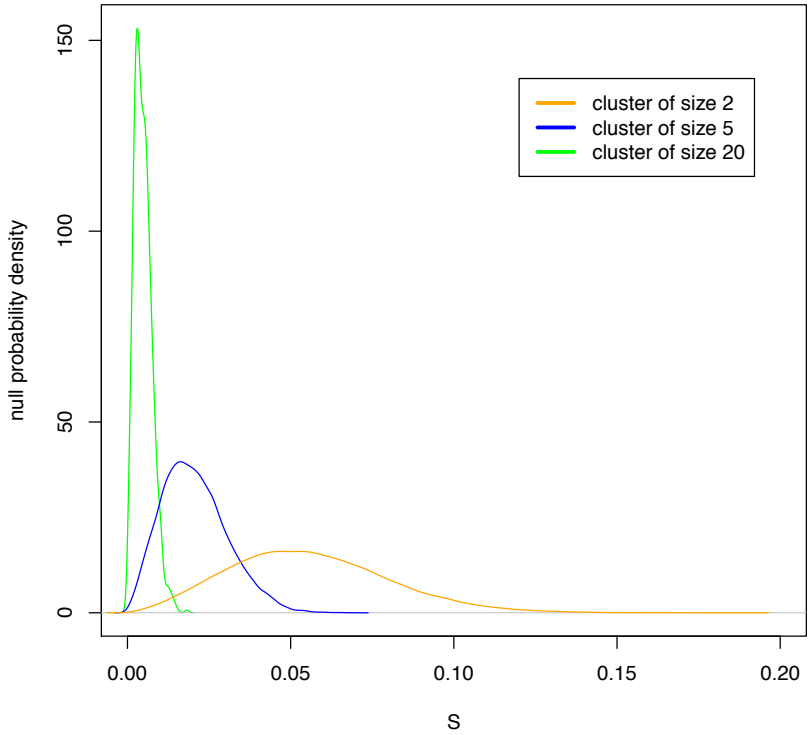
A: Simulated6

Euclidean dissimilarity – complete linkage combination



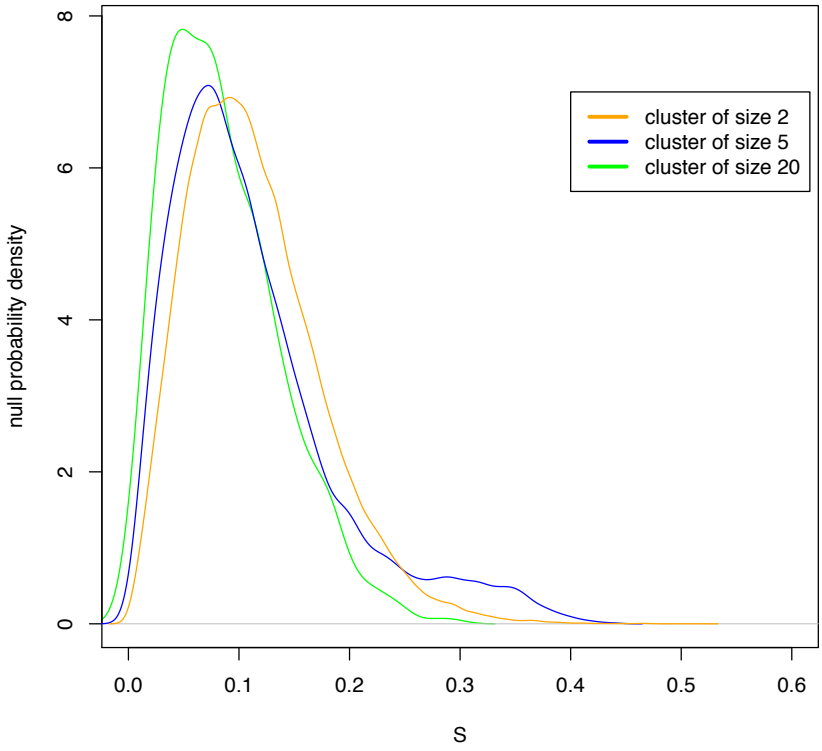
B: Simulated6

(1 - Pearson correlation) dissimilarity – average linkage combination



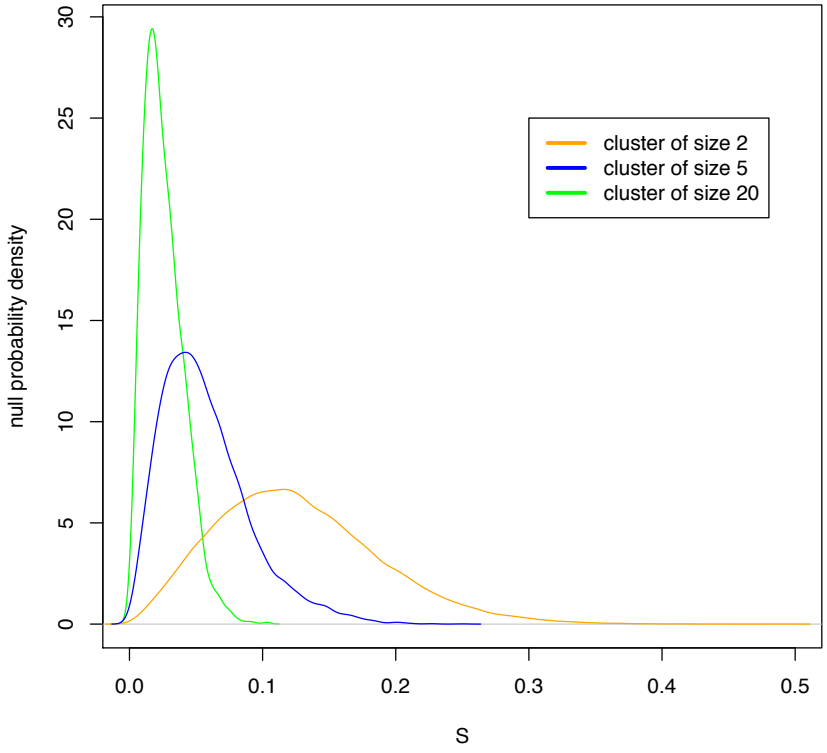
C: Leukemia

Euclidean dissimilarity – Ward linkage combination



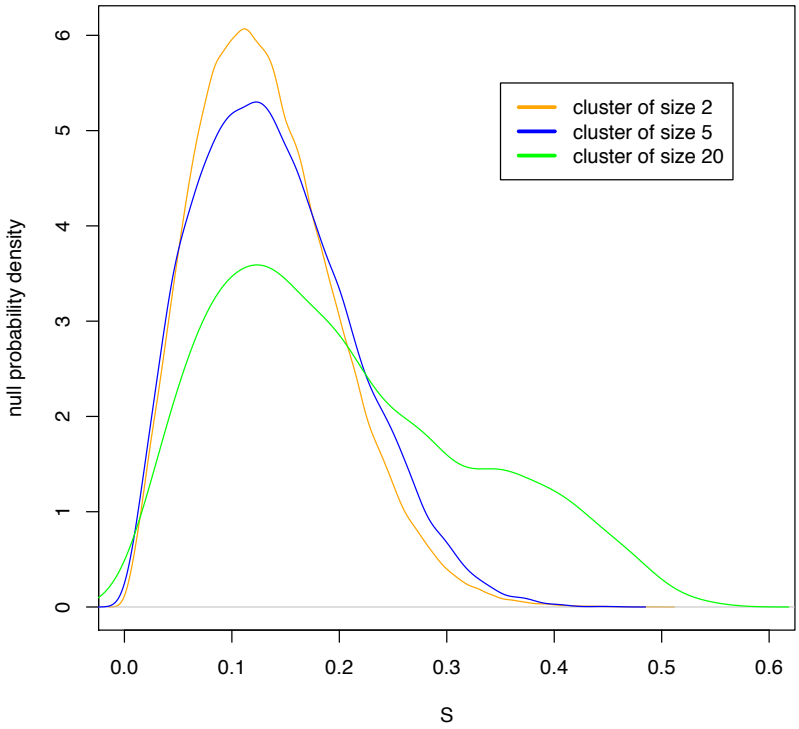
D: Leukemia

(1 - Pearson correlation) dissimilarity – average linkage combination



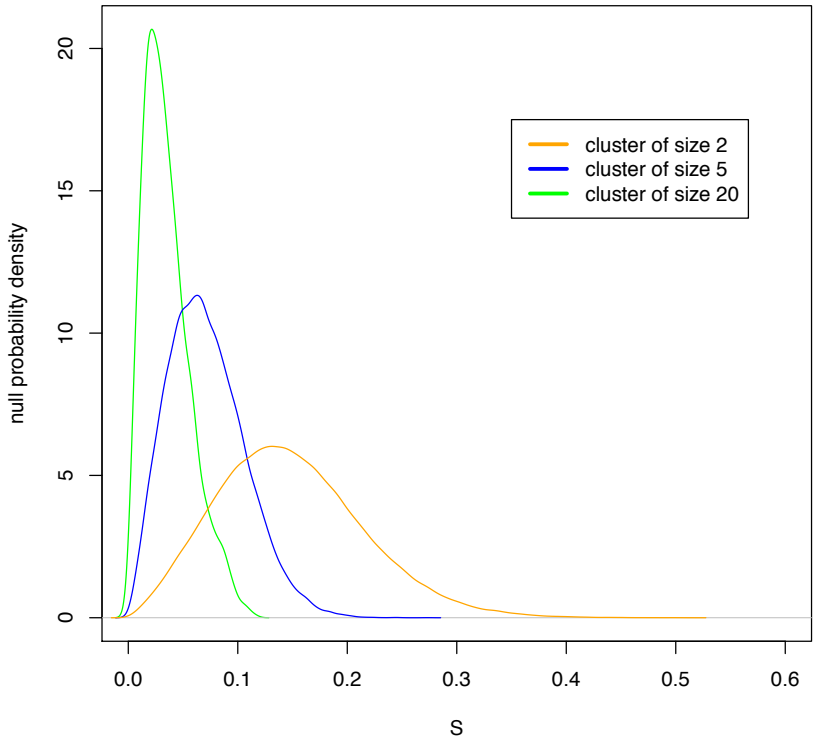
E: T10

Euclidean dissimilarity – Ward linkage combination



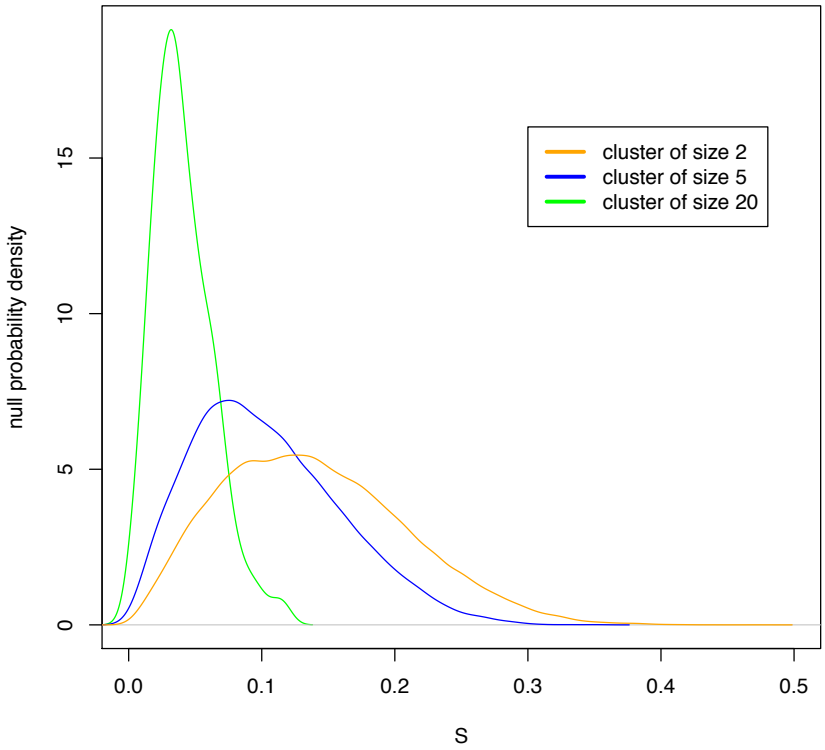
F: T10

(1 - Pearson correlation) dissimilarity – average linkage combination



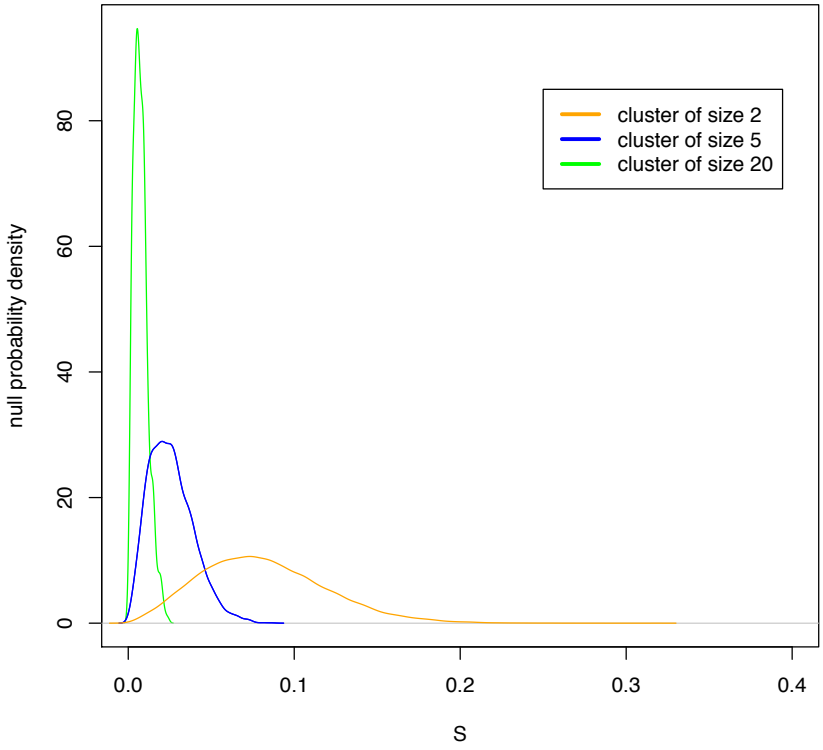
G: Organelles

(1 - Pearson correlation) dissimilarity – Ward linkage combination



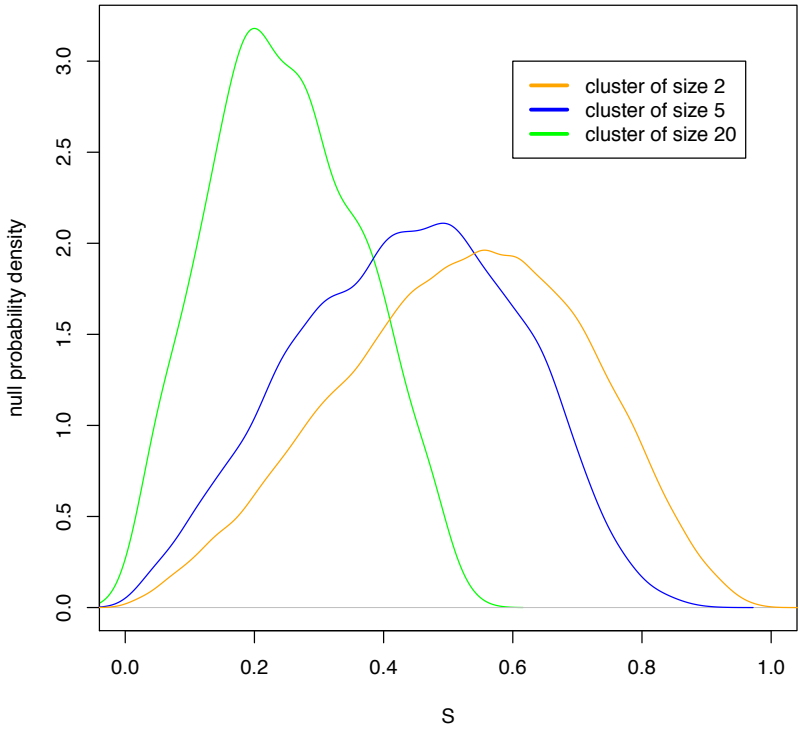
H: Organelles

(1 - Pearson correlation) dissimilarity – average linkage combination



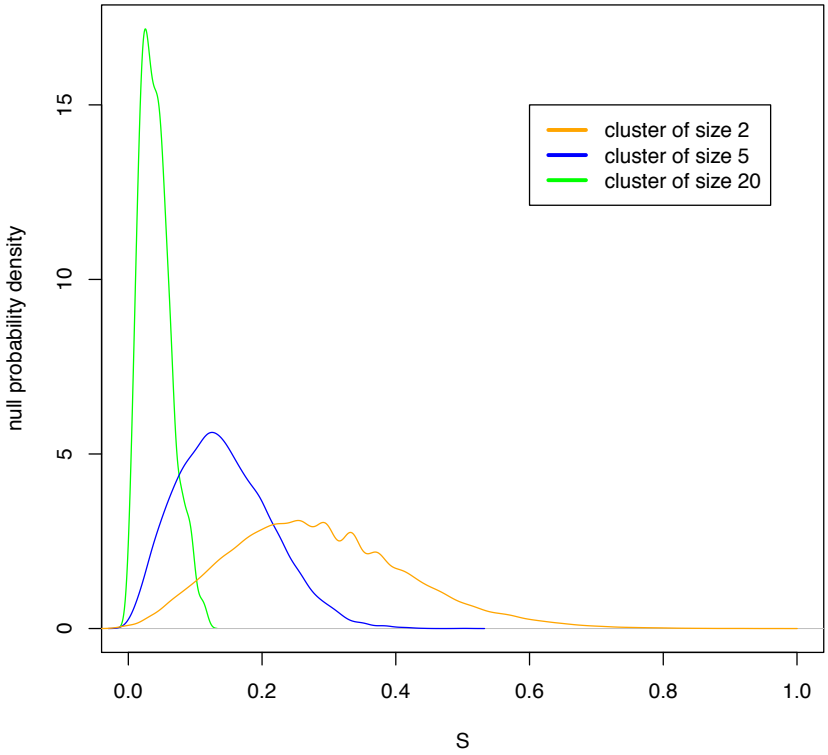
I: Chondrosarcoma

(1 - Spearman correlation) dissimilarity – Ward linkage combination



J: Chondrosarcoma

(1 - Kendall correlation) dissimilarity – average linkage combination



K: Chondrosarcoma

Manhattan dissimilarity – Ward linkage combination

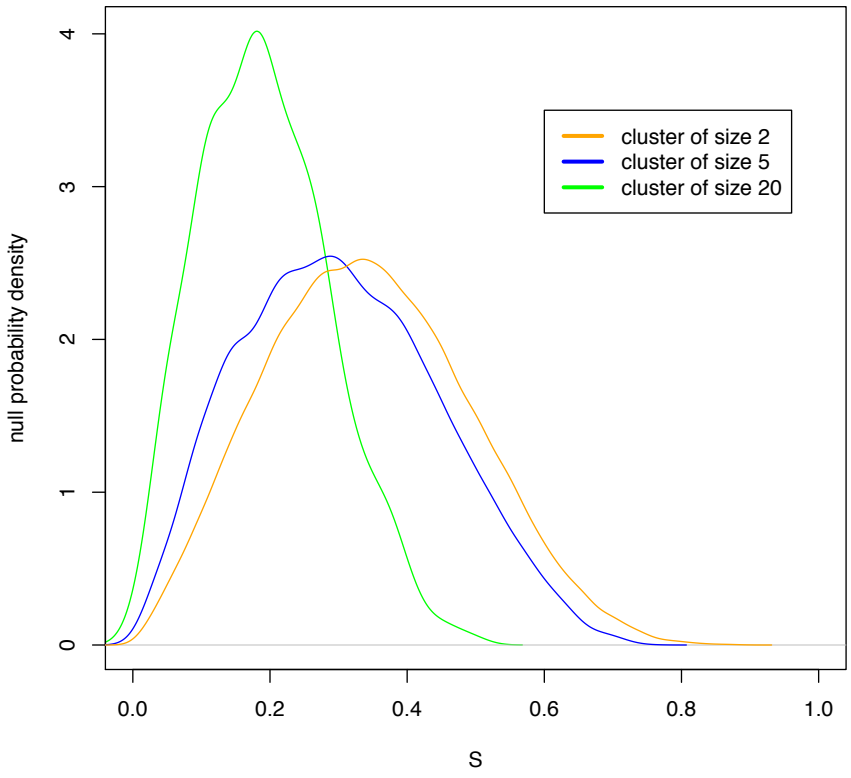
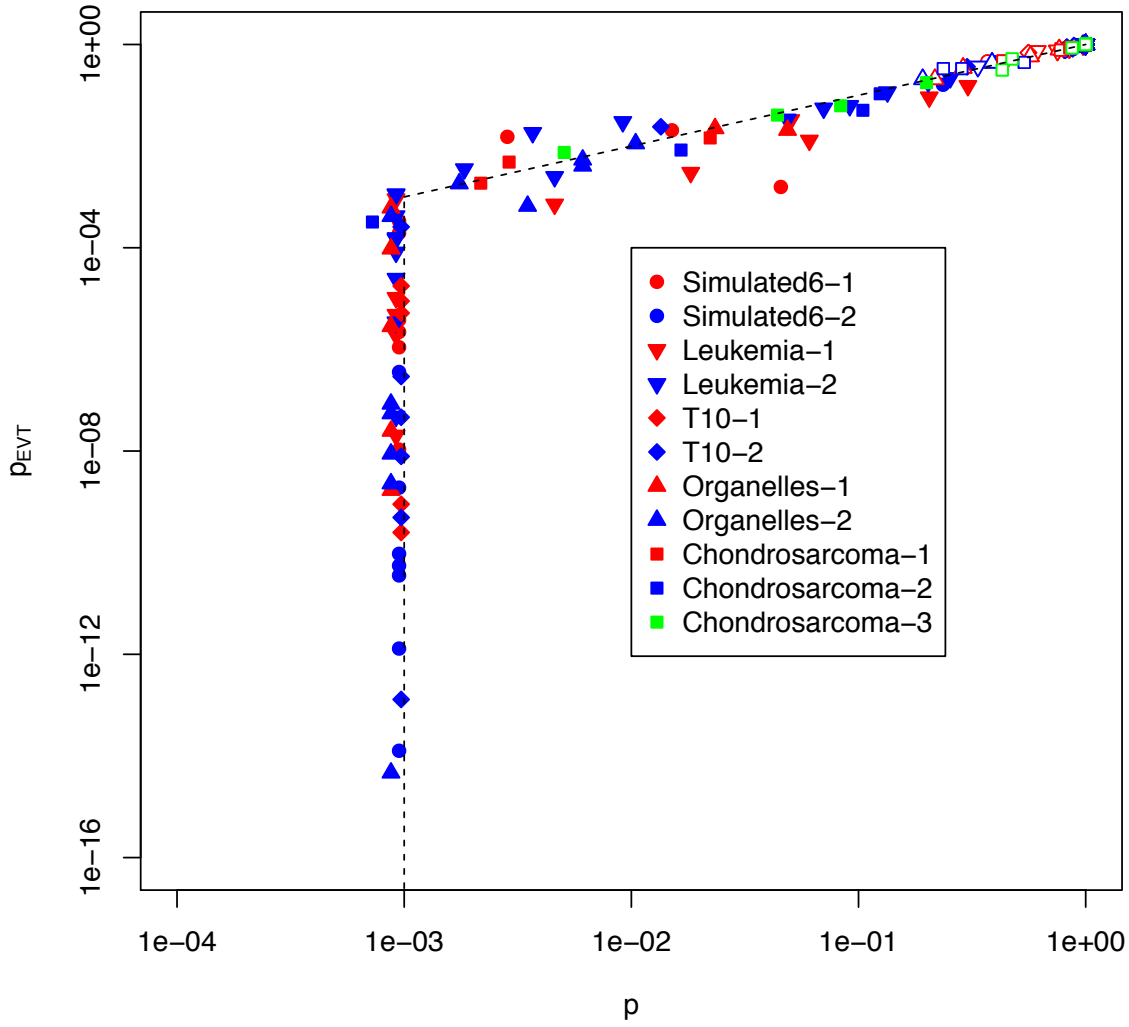


Figure S2


1e+00
1e−04
1e−08
1e−12
1e−16
pEVT
1e−04
1e−03
1e−02
1e−01
1e+00
p
Simulated6−1
Simulated6−2
Leukemia−1
Leukemia−2
T10−1
T10−2
Organelles−1
Organelles−2
Chondrosarcoma−1
Chondrosarcoma−2
Chondrosarcoma−3